

ABSTRACT

Identifying genomic sequences and oligonucleotide sequences unique to a set of organisms. Methods include obtaining genomic data characteristic of the set; formatting the data into at least one query-length sequence, each query-length sequence being of a format compatible with a similarity search engine. Searching a selected genomic database using the query and the similarity search engine. Then parsing the results of the search for those sequences showing uniqueness to the set.